# Structural Equation Modelling as an alternative to Multiple Linear Regression Models

Miriane Lucindo Zucoloto*
Universidade de São Paulo – USP, Ribeirão Preto, São Paulo, Brazil
mirianezucoloto@usp.br

João Maroco
ISPA-IU- Instituto Universitário de Ciências Psicológicas, Sociais e da Vida, Lisbon, Portugal
jpmaroco@ispa.pt

Edson Zangiacomi Martinez
Universidade de São Paulo – USP, Ribeirão Preto, São Paulo, Brazil
edson@usp.br

Juliana Alvares Duarte Bonini Campos
Universidade Estadual Paulista - UNESP, Araraquara, São Paulo, Brazil
jucampos@fcfar.unesp.br

## Abstract

The structural equation modelling (SEM) is an extension of generalized linear models that considers a causal relationship model between variables with measurement error. This technique has been applied in studies with multiple dependents and independents variables, using latent and manifest variables, with simultaneous relationship of dependence and in highly complex models, being the analyzed model based on a theoretical framework established a priori. Thus, this study was conducted in order to present the technique of structural equation modelling as an alternative to the more common analysis, such as Multiple Linear Regression. We compared two analysis (Strctural Equation Modelling and Multiple Linear Regression) in order to evaluate the better model to assess the contribution of the pain severity, age and gender on the Oral Health Impact Profile-OHIP (dependent variable) in a sample of 1,007 Brazilian dental patients. The Oral Health Impact Profile (OHIP-14) and the Multidimensional Pain Inventory (MPI) were used. The sociodemographic and clinical variables collected were gender and age. We firstly performed a Multiple Linear Regression (MLR), considering the mean scores of the OHIP and Pain severity (MPI) as manifest variables. The existence of outliers was assessed by the square Mahalanobis distance ($D^2$) and the normality (univariate and multivariate) of the variables was evaluated by shape measures (Skewness and Kurtosis). The effects were considered statistically significant when $p < 0.05$. In a second moment the data were included in a Structural Equation Model (SEM), considering the factor "OHIP" as the central construct. The fit of the model was first analysed by the evaluation of the goodness of fit indices ($\chi^2/df \leq 3.0$, CFI and GFI$\geq$0.90 and RMSEA<0.10) and the contribution of the independent variables was based on statistical significance of causal paths ($\beta$), estimated by the z test, considering a significance level of 5%. As result of the MLR analysis, only the gender was statistically significant and the total explained variance of the model was 2.0%. For the SEM analysis, we observed adequate fit of the model to the data ($\chi^2/df=1.59$; CFI=0.99; GFI=0.98; RMSEA=0.04) and the independent variables contributed 20.0% to the variability of the central construct. Thus, the SEM can be a realistic alternative to reflect the complexity and multidimensionality present in certain theoretical discussions, providing more accurate and reliable results than usual techniques, such as multiple linear regression.

**Keywords:** statistical analysis; structural equation modelling; multiple linear regression; latent variable.